A Non-Bayesian Model for Tree Crown Extraction using Marked Point Processes

Guillaume Perrin, Xavier Descombes, Josiane Zerubia

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Non-Bayesian Model for Tree Crown Extraction using Marked Point Processes

Guillaume Perrin — Xavier Descombes — Josiane Zerubia

N° 5846

February 2006

Thème COG

ISSN 0249-6399 ISRN INRIA/RR--5846--FR+ENG

A Non-Bayesian Model for Tree Crown Extraction using Marked Point Processes

Guillaume Perrin , Xavier Descombes , Josiane Zerubia

Thème COG — Systèmes cognitifs
Projet Ariana

Rapport de recherche n° 5846 — February 2006 — 34 pages

Abstract: High resolution aerial and satellite images of forests have a key role to play in natural resource management. As they enable forestry managers to study forests at the scale of trees, it is now possible to get a more accurate evaluation of the resources. Automatic algorithms are needed in that prospect to assist human operators in the exploitation of these data. In this paper, we present a stochastic geometry approach to extract 2D and 3D parameters of the trees, by modelling the stands as some realizations of a marked point process of ellipses or ellipsoids, whose points are the locations of the trees and marks their geometric features. As a result we obtain the number of stems, their position, and their size. This approach yields an energy minimization problem, where the energy embeds a regularization term (prior density), which introduces some interactions between the objects, and a data term, which links the objects to the features to be extracted, in 2D and 3D. Results are shown on Colour Infrared aerial images provided by the French National Forest Inventory (IFN)

Key-words: Marked point processes, RJMCMC, simulated annealing, automatic feature extraction, forests, tree crowns, Colour Infrared aerial images.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 —Télécopie : +33 4 92 38 77 65

Modèle non bayésien appliqué à l'extraction de houppiers par processus ponctuels marqués

Résumé : Dans ce rapport de recherche, notre but est d'extraire les houppiers à partir d'images aériennes de forêts à l'aide de processus ponctuels marqués d'ellipses ou d'ellipsoïdes. Notre approche consiste, en effet, à modéliser les données comme des réalisations de tels processus. Une fois l'objet géométrique de référence choisi, nous échantillonnons le processus objet défini par une densité grâce à un algorithme MCMC à sauts réversibles, optimisé par un recuit simulé afin d'extraire la meilleure configuration d'objets, qui nous donne l'extraction recherchée. Nous obtenons ainsi le nombre des arbres, leur localisation et leur taille. Nous présentons, dans ce rapport, un modèle 2D et un modèle 3D pour extraire des statistiques forestières. Ceux-ci sont testés sur des images aériennes infrarouge couleur très haute résolution fournies par l'Inventaire Forestier National (IFN).

Mots-clés : Processus ponctuels marqués, RJMCMC, recuit simulé, extraction automatique d'attributs, forêts, houppiers, images aériennes infrarouge couleur.

Contents

1 Introduction

The French National Forest Inventory (IFN) carries out the continuous survey of the forest resources in France. Firstly, the inventory focuses on the estimation of the surfaces pertaining to the main categories of land cover, and on the estimation of the available wood resources in production forests : volumes and yields by species, stand types and product categories. Another objective complies with environmental concerns. The method implemented by IFN relies on an extensive use of Colour Infrared (CIR) aerial photographs : each French "département" is the target of a complete aerial cover, renewed approximately every 10 years. The IFN photographic library, built from the beginning of the 1960s up to now, hosts more than 400,000 pictures representing almost 3.5 times the total area of France [12]. But this method is currently evolving. IFN tends to use the Orthophoto Database of the French National Geographic Institute (IGN) made with their 4 band digital camera (color and near infrared) and to explore 10 % of each French region each year (instead of one every 10 years). This change has been conducted in order to get annualy some representative statistics on forest resources in France, and to be able to react more rapidely to exceptional events (such as storm damages, fires, ...).

The aerial photographs, once scanned at a resolution of 50cm/pixel, represent precious data for the image processing community. The aim is to design some algorithms which would give automatically or semi-automatically some resource parameters by processing the image, and extracting its components (trees, stands, ...). Moreover, the color, the shape and the texture information of the trees in a plantation or a forest could be used in the algorithm to infer the tree species. Then, statistics such as the number of trees, their size or the density of the stands could be obtained and give some useful information at the scale of the tree which are not conceivable without any help of image processing (it would take too much time for an operator to get these statistics, and also be very costly). For instance, the minimum mapped surface by IFN is 2.25 ha in France, while smaller surfaces like alignments or isolated trees are only statistically assessed via explorations and data measurements on the ground. To be able to respond to increasing (private or public) demands, the forest inventories need such automatic algorithms.

Different algorithms for segmenting individual tree crowns without information on the Digital Elevation Model have been proposed over the past few years. For CIR images, some tools are based on a pixel-based method and give the delineation of the tree crowns, some examples are the valley following algorithm [7, 11], the

spatial clustering [4], or the region growing model [6]. Once the delineation process is finished, a classification can be foreseen using texture parameters [8] or neural networks [18]. Other tools use an object-based method, by modelling a synthetic tree crown template to find the tree top positions [10, 19]. Finally, there are some mathematical morphology methods [2], and many other ones developed for different kinds of data like LIDAR [1].

In this research report, we propose to use marked point processes [3, 21] to extract the tree crowns. They enable to model an unknown number of geometrical objects in a scene within a stochastic framework. They are increasingly exploited in image processing because they can manage to deal with both some prior information about the interaction between the features to be extracted (alignment, overlapping, ...), and some data information to fit them in the image. See for instance [5, 9, 13, 20]. Our main contribution since [16] is to consider a non Bayesian model for the energy, which enables us to get interesting results for a much wider class of images, including isolated trees, and which provides information about the height of the tree when drop shadows are distinguishable. Note that marked point processes have been used in other forestry applications such as tree crown extraction in LIDAR images [1], or statistics inference [14], once the localization of the trees is obtained.

In the first part of this report, we describe our 2D and 3D models and explain why we introduced a non Bayesian energy, starting from unconvincing results obtained with the Bayesian model [17]. Then, in the second part we test these new models on CIR images provided by IFN and comment them. Finally, we conclude and give some information about future work.

2 Forestry images at the object level

2.1 Data

The French Forest Inventory hosts a lot of CIR aerial photographs of forests which cover the whole territory. The scale of the photographs which, from one county to the other, can vary from 1/15,000 to 1/25,000, is generally around 1/20,000. The wavelengths of the three bands are between 520nm and 900nm approximately. The photographs are taken from June to September. This choice is justified by generally favourable weather conditions in summer in France, and the period of full chlorophyll activity for all plant species. In CIR images, healthy vegetation presents some peaks in the near infrared wavelength band. Moreover, the infrared reflectance changes between different species and at different periods of the year. That is why the information given by the radiometry is useful to detect and classify the trees (see [12] for more details).

Some of the photographs were scanned at a resolution of 50cm/pixel in the image. In the different models for the data energy, we assume that the ground is flat, and that the extracted trees are close to the Nadir point, in order not to take into account the induced deformation of the objects. If not specified, the computations involving radiometry are done on the near infrared band.

2.2 Objects of interest

The 2D model, used to extract tree crowns from plantations and dense forest areas, consists of a marked point process of ellipses. The associated set space S_2 is :

$$S_2 = \mathcal{P} \times \mathcal{K} = [0, X_M] \times [0, Y_M] \times [a_m, a_M] \times [b_m, b_M] \times [0, \frac{\pi}{2}[$$

where X_M and Y_M are respectively the width and the length of the image $\mathcal{I}$, (a_m, a_M) and (b_m, b_M) are the two semi-axes, and $\theta \in [0, \frac{\pi}{2}[$ the orientation of the objects (see Fig. (1), lefthandside).

The 3D model, used to extract isolated tree crowns, considers the height of the vertical trees and deals with ellipsoids. The associated state space $S_3 = S_2 \times [h_m, h_M]$ where (h_m, h_M) are the minimum and the maximum semi-height axis of the ellipsoid. Moreover, in order not to add a new unrelevant parameter, we consider that the size of the shank s is fixed (see Fig. (1), righthandside).

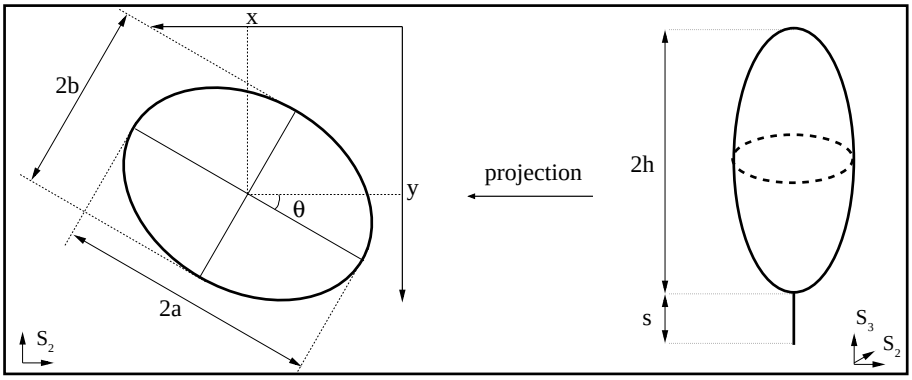

Figure 1: Left : position and marks of an ellipse of S_2. Right : additional mark for an ellipsoid of S_3.

3 From a Bayesian to a non Bayesian framework

We model the forestry images as composed of trees whose positions and attributes are some realization of a marked point process X (see [3, 21] for more mathematical details on point processes, and [5, 20] for some applications in image processing). X is also a random variable whose realizations are random configurations of objects belonging to the state space $S = \mathcal{P} \times \mathcal{K}$. Each object of S is defined by its position in $\mathcal{P}$ and the set of its marks, or attributes, in $\mathcal{K}$. The objects of the process, which account for the trees, are typically some geometric features (ellipses or ellipsoids introduced previously). We note $\mathbf{\Psi}$ the space of all configurations of a finite number of objects.

We define the probability distribution $\mathcal{P}_X(.)$ of the stochastic process as being uniformly continuous with respect to the Poisson measure $\mu(.)$ of intensity $\lambda(.)$ on S. This intensity can be homogeneous or non homogeneous. For instance, we can make the most of the infrared sensitivity of the vegetation by using the Normalized Difference Vegetation Index (NDVI) to ponder the intensity $\lambda(.)$:

$$NDVI = \frac{Infrared - Red}{Infrared + Red} \tag{1}$$

The stochastic process becomes inhomogeneous by giving more weight to some objects located in high NDVI areas, which is what we want because the trees will be located in these areas (see Fig. (2)). $\mathcal{P}_X(.)$ can be written as :

$$\mathcal{P}_X(d\mathbf{x}) = f(\mathbf{x})\mu(d\mathbf{x}) \tag{2}$$

where $f(.)$, the density of the marked point process, is detailed in the sequel. We find the seeked tree crown extration (which maximizes this density) by using an Adaptive Simulated Annealing (ASA) and a Reversible Jump Markov Chain Monte Carlo algorithm (see [15]).

In the section below, we show why we decided to develop a non Bayesian approach instead of the Bayesian one, and then we detail the new model.

3.1 The limits of the Bayesian model

Under some conditions, it is possible to write the density of the process $f(.)$ using Bayes rule :

$$f(\mathbf{x}) = f(\mathbf{x}|\mathcal{I}) = \frac{f_p(\mathbf{x})\mathcal{L}(\mathcal{I}|\mathbf{x})}{f(\mathcal{I})} \propto f_p(\mathbf{x})\mathcal{L}(\mathcal{I}|\mathbf{x}). \tag{3}$$

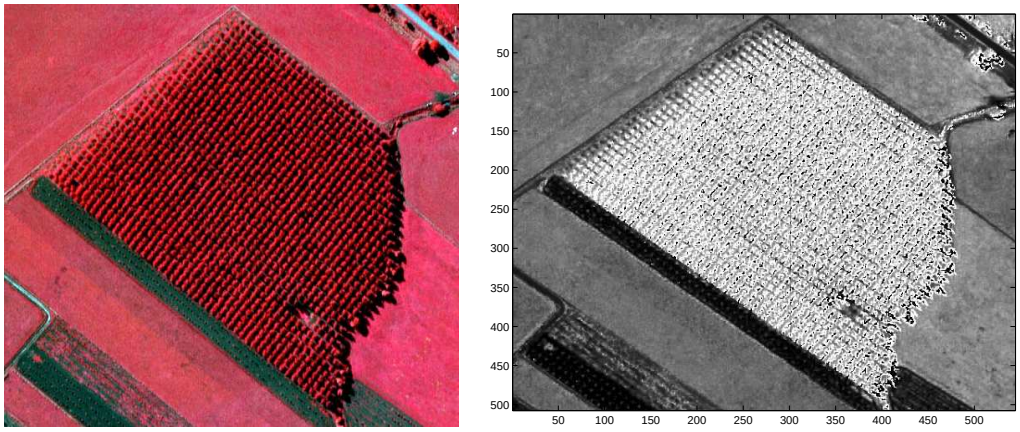

Figure 2: Left : original image © IFN. Right : NDVI : intensity $\lambda(.)$ of the reference Poisson process.

Basically, it requires first to define a likelihood $\mathcal{L}(\mathcal{I}|\mathbf{x})$, which represents the probability for the data $\mathcal{I}$ knowing the configuration $\mathbf{x}$. In our application, this requires to be able to model the grey levels of the image both inside and outside the objects of the configuration. We have to be able to know the pixel distribution outside these objects which is not so obvious in remote sensing images (containing a great variety of types of objects : crops, buildings, lakes ...). That is why a pre-segmentation of the image is needed before, in order to work directly on the plantation itself.

To define the likelihood, we consider a Gaussian mixture which models the data. Each pixel in the image is associated to one of the two Gaussian classes :

- $\mathcal{C}_i = \mathcal{N}(m_i, \sigma_i)$ for the pixels inside at least one of the objects of the configuration,
- $\mathcal{C}_o = \mathcal{N}(m_o, \sigma_o)$ for the pixels outside.

However, to take into account the variation of the pixel intensity of the tree crowns depending on their position with respect to the tree top, we can consider some adaptative Gaussian mixture (see [17]). The likelihood $\mathcal{L}(\mathcal{I}|\mathbf{x})$ is :

$$\mathcal{L}(\mathcal{I}|\mathbf{x}) = \prod_{p \in \mathcal{C}_i} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(\frac{-(y_p - \mu_i)^2}{2\sigma_i^2}\right) \prod_{p \in \mathcal{C}_o} \frac{1}{\sqrt{2\pi}\sigma_o} \exp\left(\frac{-(y_p - \mu_o)^2}{2\sigma_o^2}\right). \quad (4)$$

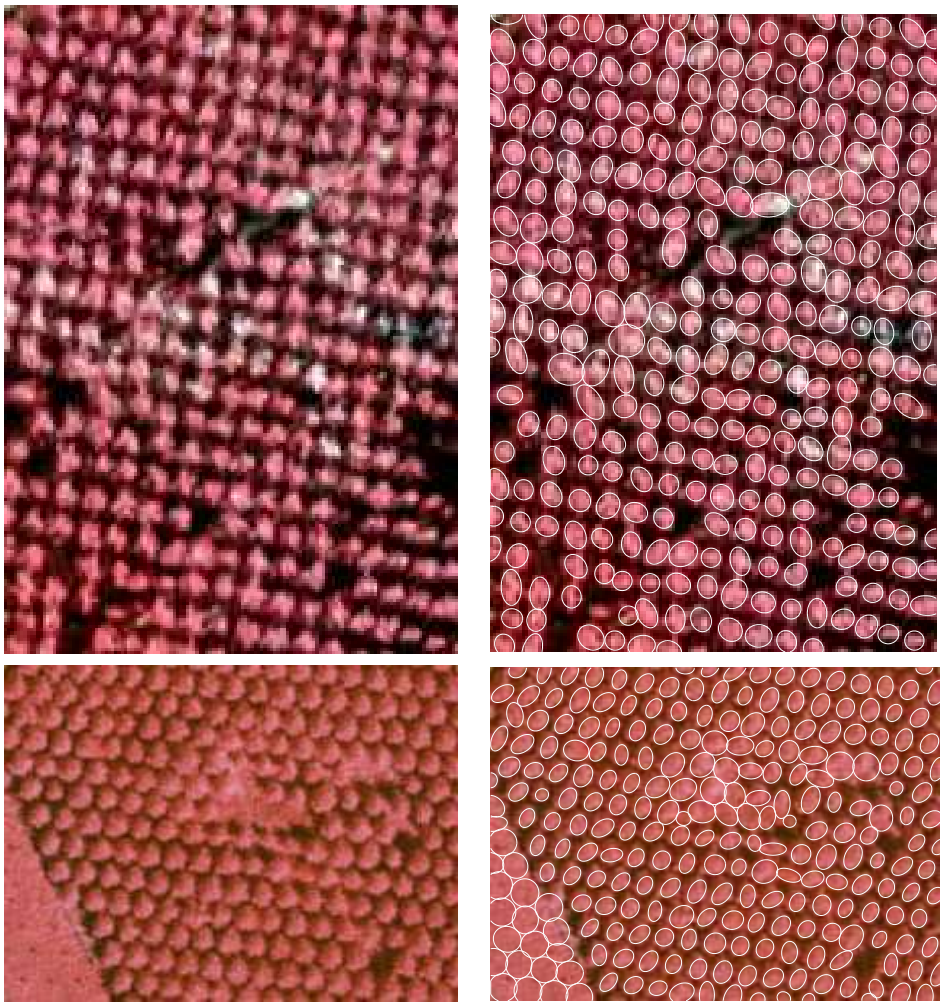

Figure 3: Some extraction results on aerial photographs using the Bayesian model. Top : original image (220×140) and the extraction result. Bottom : original image (166×224 pixels) and the extraction result. All original images © IFN, all extraction results © INRIA.

Results presented in Fig. (3) show that the Bayesian model works well (see top right) when the image contains only trees and a background which does not contain vegetation (see top left). However, when the pre-segmentation of the image is not available, the energy is minimized when the pixels of the crops belong to the tree class, that is why so many false alarms are detected in the bright part of the image (see bottom left and bottom right). Thus, we have to change the data term and define a better characterization of a tree in a forestry image.

3.2 Non Bayesian models

Using the Gibbs formulation, we can rewrite (2) as :

$$\mathcal{P}_X(d\mathbf{x}) = \frac{1}{Z} \exp\left(-U(\mathbf{x})\right) \mu(d\mathbf{x}) \tag{5}$$

where Z is a normalizing constant and $U(\mathbf{x})$ the energy of the configuration $\mathbf{x}$. The energy is the sum of the prior energy $U_p(\mathbf{x})$, which takes into account the interactions between the objects of the configuration, and the data energy $U_d(\mathbf{x})$, which links the objects to the data :

$$U(\mathbf{x}) = U_p(\mathbf{x}) + U_d(\mathbf{x})$$

The prior energy, detailed in [17], is written as :

$$U_p(\mathbf{x}) = \gamma_r R(\mathbf{x}) + \gamma_a A(\mathbf{x}) \tag{6}$$

where $\gamma_r > 0$ and $\gamma_a < 0$ are respectively some weights which ponder the repulsion between the objects of the process and the attraction between aligned objects in a plantation. Below, we detail the data term of the energy. In the non-Bayesian model, the data energy is not calculated at the pixel level anymore, but at the object level :

$$U_d(\mathbf{x}) = \gamma_d \sum_{x_i \in \mathbf{x}} U_d(x_i)$$

where $\gamma_d > 0$ and $U_d(x_i) \in [-1, 1]$. An object will be attractive and therefore favored if its data energy is negative. On the contrary, a positive data energy will yield a repulsion for the related object.

In the following, we detail two models for the data energy term. The first one is called the 2D model which deals with ellipses as objects of the process. The second one is the 3D model, and deals with ellipsoids.

3.2.1 2D model

In dense areas (we call "dense area" a stand where the drop shadows of the trees are merging together : a plantation for instance), the tree crowns can be extracted thanks to their radiometry (picks in the near infrared band) and some shadow that surrounds them. These dark pixels around the trees are used by all the existing segmentation and delineation algorithms. Thus, we define the boundary of an ellipse $\mathcal{F}(x)$ as the set of $\mathcal{P}$ contained between the given ellipse $x = (p, k)$, where $k = (a, b, \theta)$ are the marks, and a concentric one $x' = (p, k')$, with $k' = (a + \rho, b + \rho, \theta)$ (see Fig. (4)). This boundary will stand for the shadow that comes from the tree itself and from its neighbours. Then, we calculate the Bhattacharya distance $d_B(x, \mathcal{F}(x))$ between the reflectance distributions of the object and its boundary, supposed to be Gaussians, and define the data energy $U_d(x)$ as :

$$U_d(x) = \mathcal{Q}_b\left(d_B(x, \mathcal{F}(x))\right)$$

where $\mathcal{Q}_b(d_B) \in [-1, 1]$ is a quality function which gives a positive value to a small Bhattacharya distance (badly located objects) and a negative value (well located objects) otherwise.

$$\mathcal{Q}_b(d_B) = \begin{cases} 1 - \left(\frac{d_B}{d_0}\right)^{\frac{1}{3}} & \text{if } d_B < d_0 \\ \exp\left(\frac{-(d_B - d_0)}{3d_0}\right) - 1 & \text{otherwise} \end{cases}$$

In practice, we choose $\rho = 2$ and $d_0 = 30$ if not specified.

3.2.2 3D model

In mixed height stands, low density forest or isolated trees, individual tree shadows can be observed and help a lot to detect the trees. Then, our first idea to model the drop shadow was to use the ellipse process, and to change the definition of the boundary. Two simple boundary models for the drop shadow have been developped (see Fig. (5)). The first one just keeps the set of $\mathcal{P}$ of the original boundary which is around an angle α (typically the direction of the sun light). The second one translates the object in the direction α at a distance ρ and keeps the set of $\mathcal{P}$ which is outside the object. The data energy is still :

$$U_d(x) = \mathcal{Q}_a\left(d_B(x, \mathcal{O}^{\rho,\alpha}(x))\right) \tag{7}$$

where $\mathcal{O}^{\rho,\alpha}(x)$ is the set of pixels which refers to the shadow.

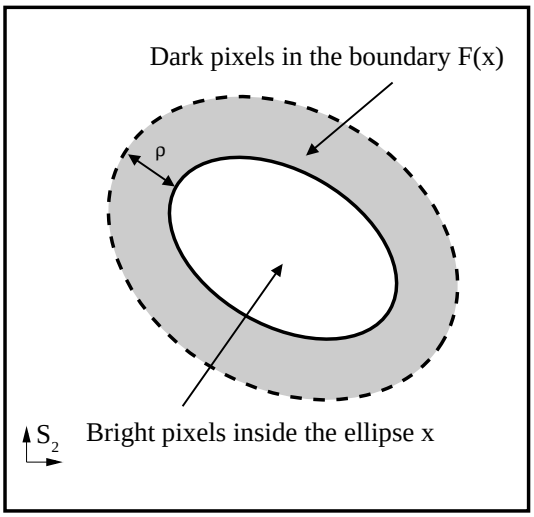

Figure 4: An ellipse and its boundary $\mathcal{F}^{\rho}(x)$.

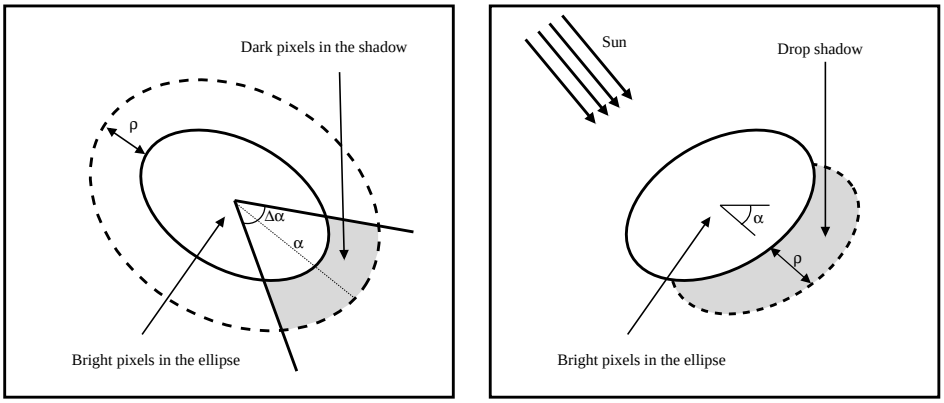

Figure 5: Modeling the drop shadow with the ellipse process by modifying the boundary. Left : the new boundary is a part of the former boundary, around an angle α. Right : the new boundary is the translated ellipse.

Then, knowing where and when the photograph has been taken, we can get the position of the sun and deduce from the shadow and the parallax the height or rela-

tive height of the trees. This new feature is useful to calculate more precisely some statistics at the scale of the stand such as the wood volume, or to retrieve some physiological parameters from the tree. However, the two simple models described previously cannot give any information about the height of the tree, because they do not fit the shadow. That is why we propose a 3D model, which deals with a 3D object process (ellipsoids) instead of the ellipse process.

The dark pixels are now searched in the tree own shadow, result of the projection of the ellipsoid on $\mathcal{P}$ in the direction of the sun light (see Fig. (6)). We define $\mathcal{S}(x, \mathbf{x})$, the shadow of an ellipsoid x belonging to $\mathbf{x}$, as the set of $\mathcal{P}$ containing this projection and excluding all the orthogonal projections of the other objects in $\mathbf{x}$. Note that if the sun cannot be located, its orientation could become a new parameter to be estimated in the algorithm. This will be tested in the next part.

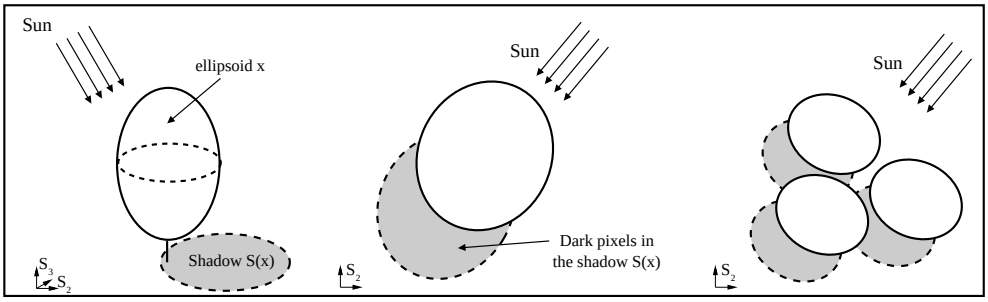

Figure 6: Left : an ellipsoid and its shadow $\mathcal{S}(x, \mathbf{x})$, 3D view. Middle : an ellipsoid and its shadow $\mathcal{S}(x, \mathbf{x})$, 2D view. Right : competition between some ellipsoids and their respective shadows.

To calculate the data energy term of an ellipsoid x, we first calculate the Bhattacharya distance $d_B(x, \mathcal{S}(x))$ between the reflectance distributions of the object and its shadow, and its quality $\mathcal{Q}_b(d_B) \in [-1, 1]$. Unfortunately, this is not sufficient to segment correctly the objects : ellipsoids could have a good d_B value without delineating correctly the border of the tree. Indeed, nothing tends to stretch the objects to the limit of the tree in the 3D model (especially in the opposite direction of the shadow), because the dark pixels which were all around the border in the 2D model, are now just on the side of the drop shadow. To solve this problem, we add a gradient term $\mathcal{Q}_g\left(G(x, \mathcal{S}(x))\right) \in [-1, 1]$ which tends to maximize the gradient G

on all the contours of both the object x and its shadow $\mathcal{S}(x)$. A texture parameter $\mathcal{Q}_t\left(T(x)\right)$ can be added to extract predefined species in a mixed stand. Finally, the data energy term is given by :

$$U_d(x) = \alpha_b \mathcal{Q}_b\left(d_B\right) + \alpha_g \mathcal{Q}_g\left(G\right) + \alpha_t \mathcal{Q}_t\left(T\right)$$

where $\alpha_b + \alpha_g + \alpha_t = 1$.

4 Results

4.1 Remarks about the simulations

Results presented in this part were obtained with a Red Hat Linux 3GHz machine. The simulations and the optimizations are managed by an Adaptive Simulated Annealing scheme embedded in a Reversible Jump Markov Chain Monte Carlo algorithm (see [15]).The perturbations of the kernel are birth and death, translation, dilation, rotation, split and merge, and birth and death in a neighbourhood for plantation images (see [17] for more information about the simulations).

Note that using the NDVI information in the definition of the intensity $\lambda(.)$ of the reference Poisson process affects the Green acceptance ratio. For example, the ratio for a simple move (translation, rotation, dilation) which proposes to transform an uniformly chosen object $u \in \mathbf{x}$ into v is now :

$$R_{NJ}(\mathbf{x}, \mathbf{y} = \mathbf{x} \setminus \{u\} \cup \{v\}) = \frac{f(\mathbf{y})\lambda(v)}{f(\mathbf{x})\lambda(u)} \tag{8}$$

If not specified, the parameters of the density in the following experiments will be $\gamma_r = 1000$, $\gamma_a = -5$ and $\gamma_d = 50$. The temperature of the Simulated Annealing is initially set to $T_0 = 10000$, and is updated every 10000 iterations with respect to an adaptive schedule (see [15]).

4.2 2D model

4.2.1 Bayesian model versus non Bayesian model

The first result we present in Fig. (7) compares a tree crown extraction performed with the Bayesian model and the non Bayesian model. The data on which we have chosen to compare the two algorithms contain some poplars, a small crop on the left, and some undetermined areas inside the plantation (possibly some very dense areas,

where it is impossible to delineate the trees even by visual inspection).

As expected, the Bayesian model classifies the bright pixels of the crop as being part of the tree crown class, and then puts some objects on this part of the image. This results in more than 30 false alarms. The non Bayesian model answers better in this zone and in the other difficult part of the image, thanks to the data energy which is attractive only if an object has some shadow that surrounds it.

4.2.2 Non Bayesian model on real data : possibilities and limits

We present in Fig. (8) some results obtained with our 2D data energy model. The first image is a zoom on a plantation, and contains approximately 300 trees. The second one represents a complete plantation containing more than 1300 trees and its neighbourhood (7 hectares). The model gives satisfactory results for these two examples.

Computation takes a couple of minutes for each result (between 2 and 10 millions of iterations). It depends highly on both the number of objects to be extracted and the size of the image. Depending on the desired accuracy or the available time, the simulation can be accelerated or slowed by changing the speed of the temperature decrease (in the adaptive scheme, see [15]). Obviously, the minimization of the energy is better for a slow decrease of the temperature, and the extraction result too. Note that the alignment favouring part of the prior term and the birth and death in the neighbourhood move in the proposition kernel are not mandatory, they just decrease the computational time (see Fig. (9) for a result without them). We can see that in this case some objects have been accepted in the top of the image because of the shadows of the border trees. A better use of the color band or a better definition of where is the shadow in the 2D model could give better results.

Moreover, we test the 2D model on a more dense plantation in Fig. (10). The problem is that in this case trees are very close to each other, especially in one direction (top-bottom) where we can hardly detect some dark pixels around them. Thus, the Bhattacharya distance between the boundary $\mathcal{F}(x)$ and the ellipse is not high enough. To solve this problem, we first try to reduce the size of the boundary, by reducing ρ from 2 to 1, but this cannot help. Likewise, dropping the threshold $d_0 = 30$ to $d_0 = 20$ accepts more objects on the one hand, but favors the proliferation of false alarms on the other hand.

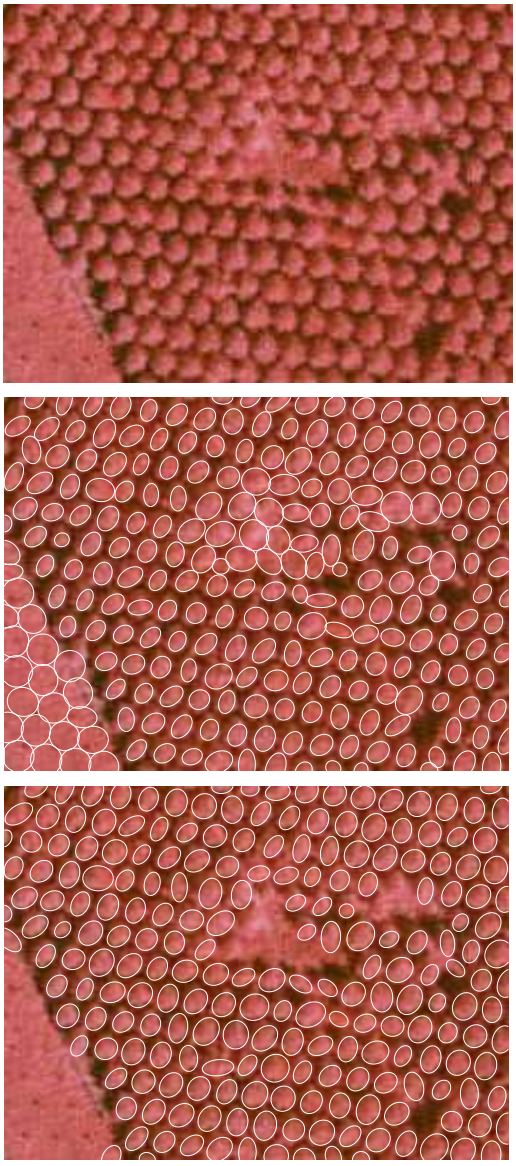

Figure 7: Top : the original data (166 × 224 pixels) ©IFN. Middle : extraction result with a Bayesian model ©INRIA. Bottom : extraction result with a non Bayesian model ©INRIA.

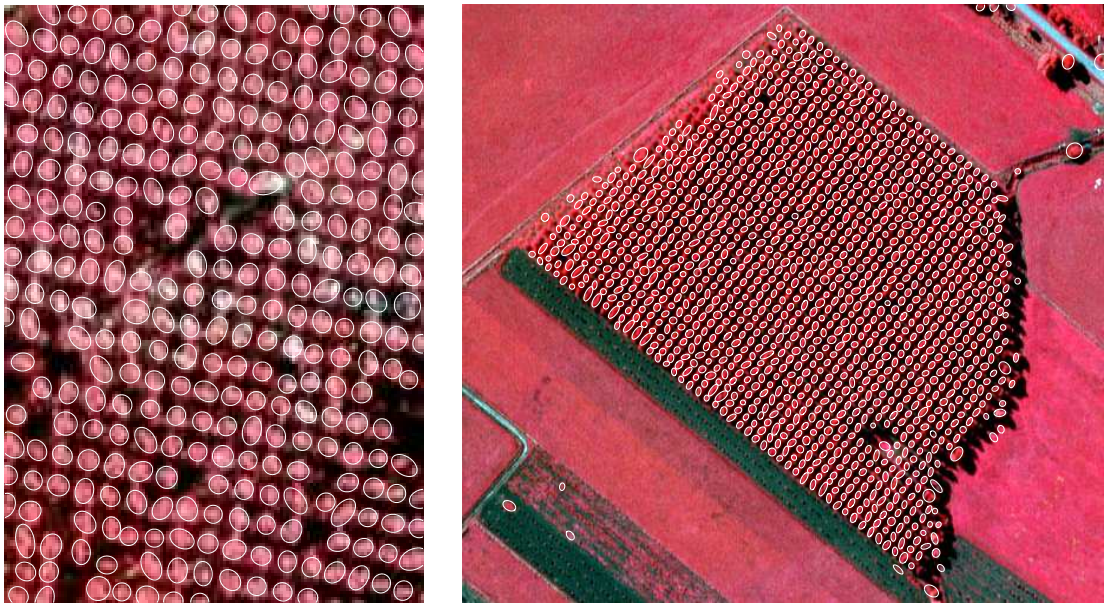

Figure 8: Left : 220×140 image of poplars. Right : 510×540 image of poplars. For both images ©INRIA

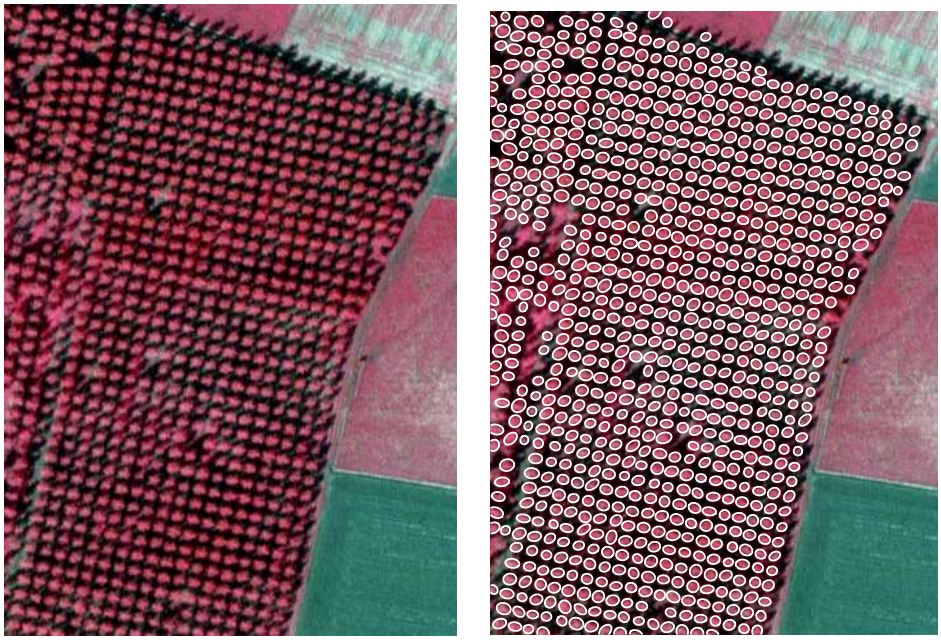

Figure 9: Left : 349×487 image of poplars ©IFN. Right : extraction result without the prior term, 3 millions of iterations. ©INRIA

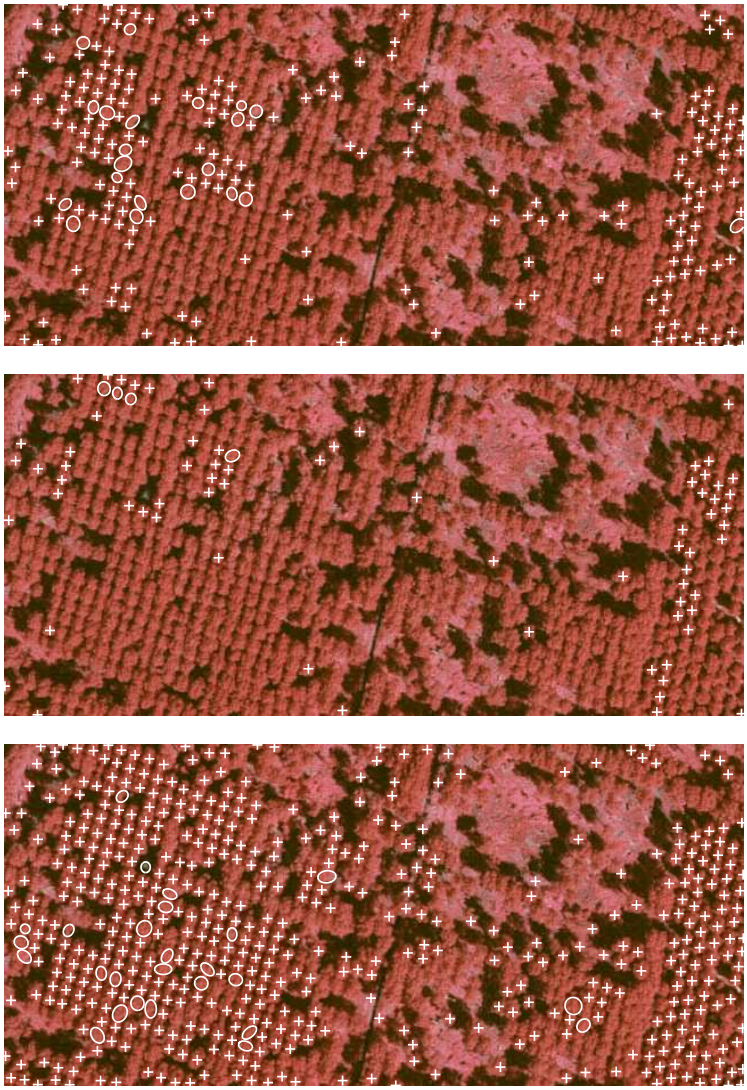

Figure 10: Top : extraction result with the 2D model and ($\rho = 2, d_0 = 30$) performed on a 300×650 pixels data (227 objects). Middle : extraction result with the 2D model and $\rho = 1, d_0 = 30$ (75 objects). Bottom : extraction result with the 2D model and $\rho = 2, d_0 = 20$ (533 objects). For all images ©INRIA. For more visibility, extracted objects with a negative data energy are represented by crosses, while extracted objects with a positive data energy (which have been accepted thanks to the prior term) are represented by ellipses.

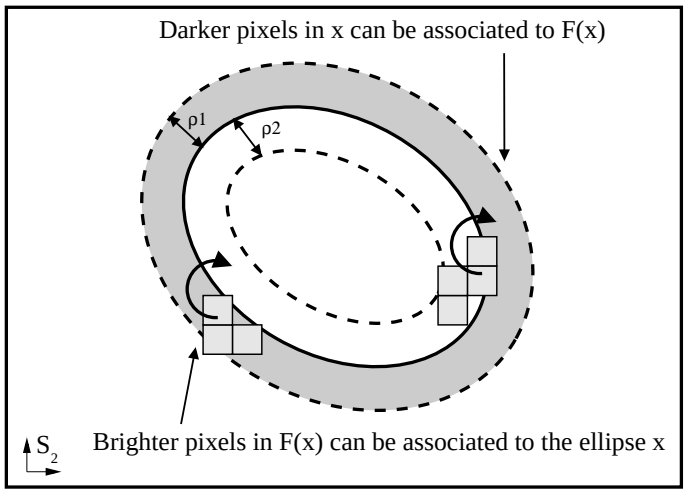

Figure 11: An ellipse and its corrected boundary $\mathcal{F}^{\rho}(x)$.

Then, we propose to adjust the definition of the data energy in order to work with more flexible objects, which could extract tree crowns in high density stands. The idea is to enable some pixels in the boundary (up to a fixed percentage p_{2D}) to be included in the ellipse during the statistical tests (Bhattacharya distance) and some pixels in the ellipse (close to the border) to be included in the boundary at the same time. This is detailed in Fig. (11).

This correction of the 2D model data term gives more flexibility to the geometry of the extracted object and reacts better for high density stands. Another idea is to increase the weight of the prior term to force the acceptance of objects even if the data term is positive. Both solutions work well, as shown in Fig. (12) on a smaller image, but the first one is preferable because it still gives a sense to the data energy (only good objects are accepted), while the second one involves a calibration of the prior energy with respect to the data energy. Moreover, removing the prior energy attraction in the case of non-plantation but very dense area should still give good results, which is only possible with the correction of the 2D model.

In Fig. (13), we present the final result obtained with the correction of the data term on the big image. We can see that more objects are extracted (713 instead of

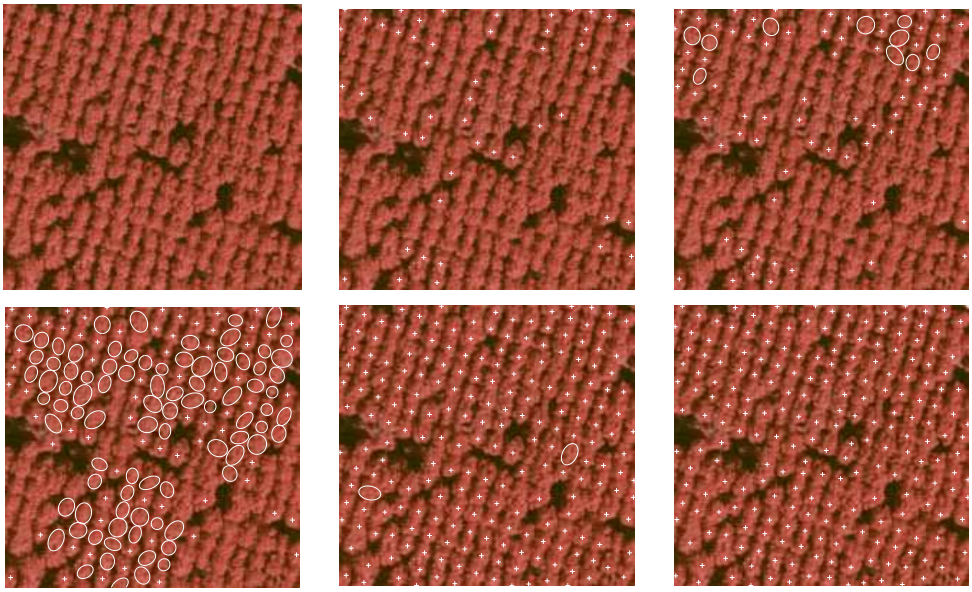

Figure 12: Top left : original image ©IFN. Top middle : 2D model without correction, $\left|\frac{\gamma_p}{\gamma_d}\right| = 0$ (49 objets). Top right : 2D model without correction, $\left|\frac{\gamma_p}{\gamma_d}\right| = 0.1$ (79 objets). Bottom left : 2D model without correction, $\left|\frac{\gamma_p}{\gamma_d}\right| = 0.2$ (159 objets). Bottom middle : 2D model with correction, $\left|\frac{\gamma_p}{\gamma_d}\right| = 0.1$ and $p_{2D} = 10\%$ (214 objets). Bottom right : 2D model with correction, $\left|\frac{\gamma_p}{\gamma_d}\right| = 0$ and $p_{2D} = 15\%$ (217 objets). For all result images ©INRIA. For more visibility, extracted objects with a negative data energy are represented by crosses, while extracted objects with a positive data energy (which have been accepted thanks to the prior term) are represented by ellipses.

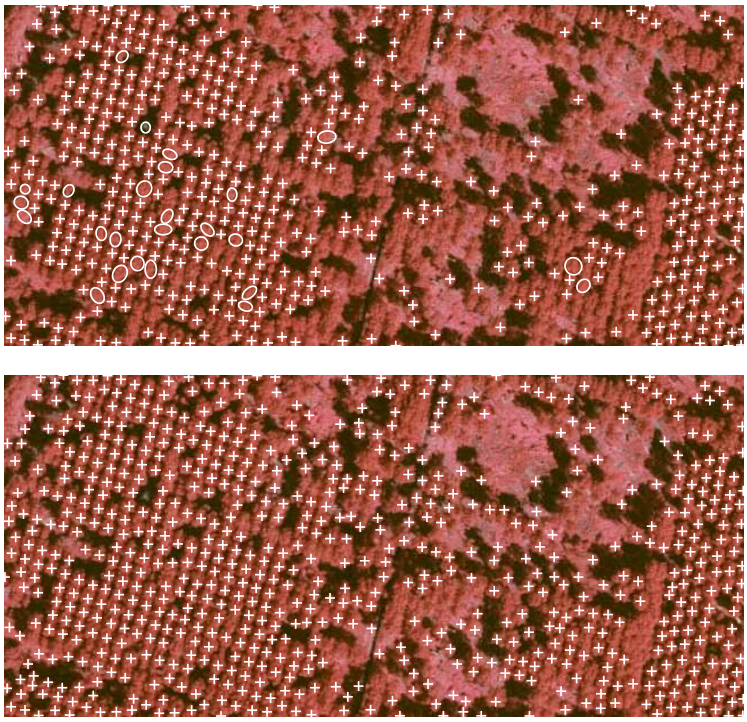

Figure 13: Top : extraction result with the 2D model and $(\rho = 2, d_0 = 20)$: 227 objects. Bottom : extraction result with the corrected 2D model : 713 objects. For both images ©INRIA. For more visibility, extracted objects with a negative data energy are represented by crosses, while extracted objects with a positive data energy (which have been accepted thanks to the prior term) are represented by ellipses.

227), but there are still some false alarms which would need to be taken into account in a future work.

4.3 3D model

In this section we present some results obtained with the proposed 3D data energy model. To the best of our knowledge, this 3D model is the only one proposed in the literature to be able to extract 3D tree crown parameters in CIR images.

First of all, we test our 3D model on a synthetic image containing 13 objects. The position of the sun has been considered as a parameter to be optimized in the algorithm : a move in the proposition kernel proposes to change the orientation of the sun. We can see that the results are convincing (see Fig. (14)).

Then, we test our algorithm on real data, but this time we fix the position of the sun (because sometimes this position was blocked at $\theta_S + \pi$ instead of the good value θ_S during the optimization). This is not a big deal because, in practice, these orientation parameters can be obtained or at least approximated from the knowledge of when and where the photograph has been taken.

The first image is a 600×550 pixel image, and contains a few isolated trees with their drop shadows. More than 750 trees have been detected (see Fig. (15)). However, some trees have not been extracted in very dense areas, and some false alarms stand at the left handside of the green crop, because of the low infrared reflection of this crop, which is considered as a shadow. Then, as noticed previously for Fig. (8), it would definitely be interesting to make a better use of the 3 bands (near infrared and colour) in Fig. (16), when using the information of the height of the trees, which is given in the 3D model.

Then, the 3D model is used on the border trees of the plantation extracted in Fig. (8) : in a post processing, we select those trees at the border of the plantation and apply the 3D model to obtain their height thanks to their drop shadows (see Fig. (17)). From this result, we can get an estimation of the wood volume, by supposing that the height is homogeneous inside the plantation. We use the AMAP Orchestra Software to reconstruct the plantation in 3D, given the size of the trees. We choose some 20 year old poplars, and an arbitrary position of the sun.

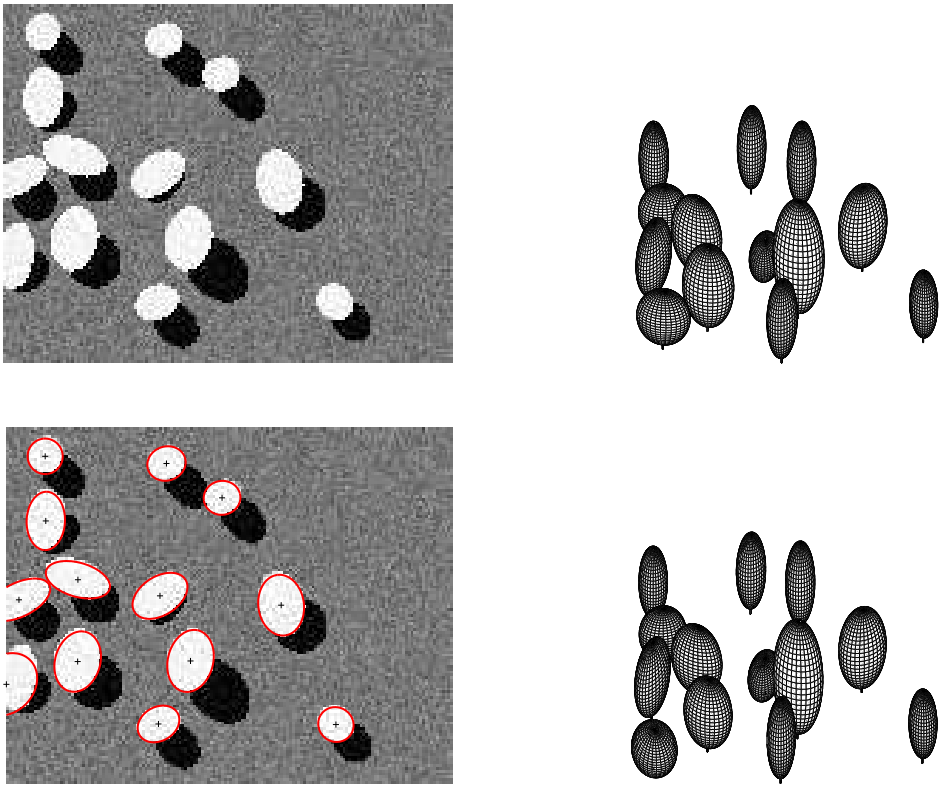

Figure 14: Top : the original synthetic image and its 3D reconstruction. Bottom : the extraction result and the 3D reconstruction of this extraction. The orientation of the sun was a parameter to be optimized in this test.

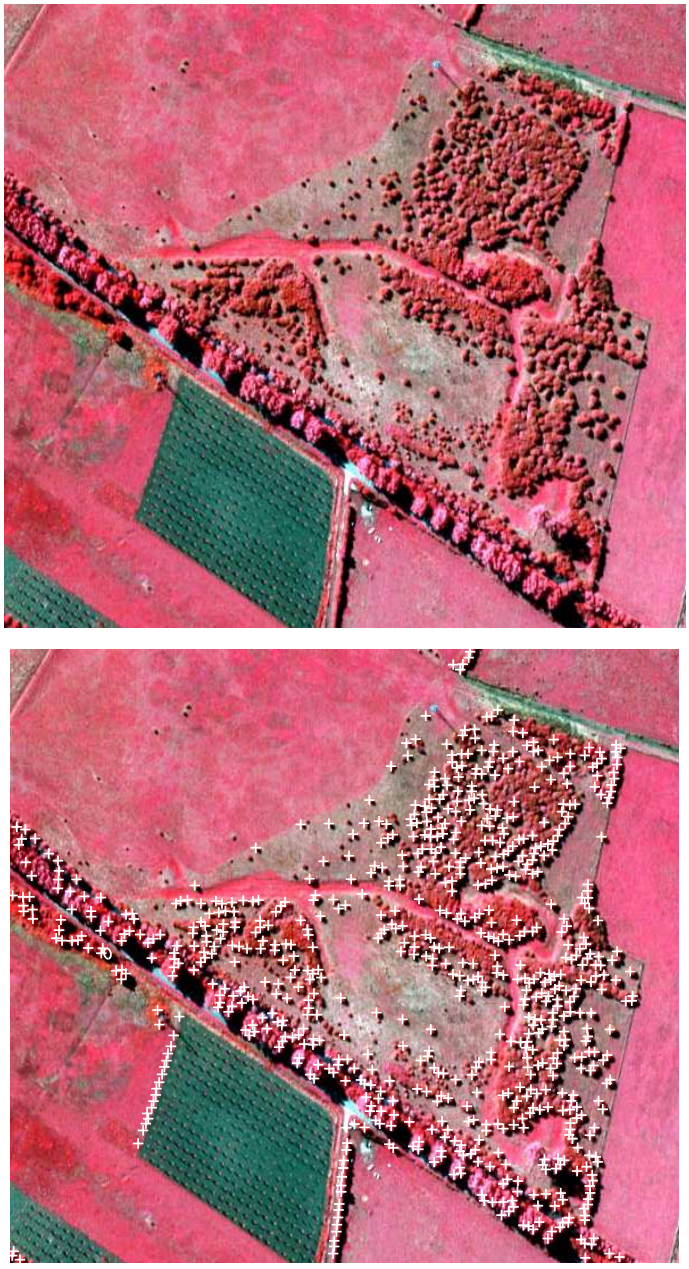

Figure 15: Left : 600×550 image of isolated trees ©IFN. Right : extraction result, 10 millions of iterations, ©INRIA.

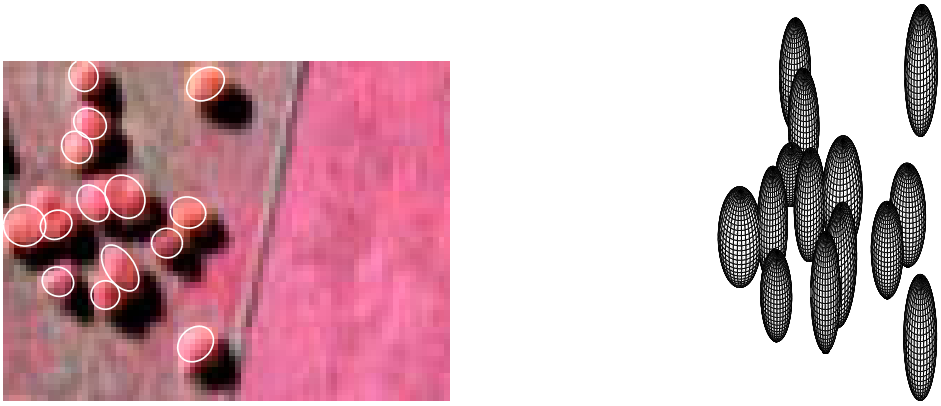

Figure 16: Left : image with 14 isolated trees ©INRIA. Right : 3D reconstruction (the position of the sun has been manually selected).

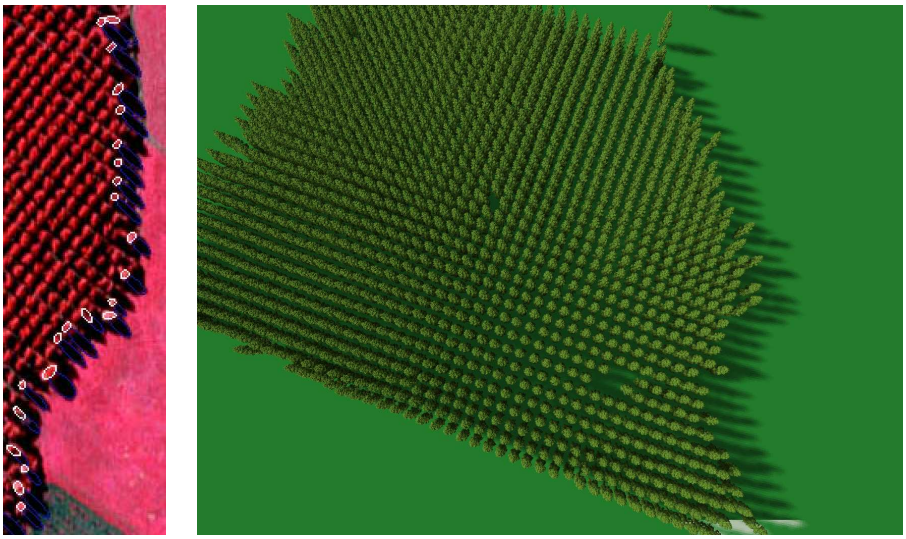

Figure 17: Left : extraction result with the 3D model on the border trees ©INRIA. Right : 3D reconstruction of the plantation with AMAP Orchestra Software.

Finally, a result is presented on a mixed height stand in Fig. (18), where we aim at detecting the big trees in the "cauliflower" structure. Indeed, the percentage of cover of such a big tree structure (timber forest) is used in the classification of the stand in the forestry maps. On this image, the 3D model is unable to count precisely the timber forest trees : for instance some big trees are extracted by several objects. In the future, we should adapt the energy of the 3D model, by including a volume parameter which favors big objects for example, in order to avoid these errors. However, by taking the area covered by the objects of the process, we have an estimation of the timber forest cover percentage ($\simeq 15\%$).

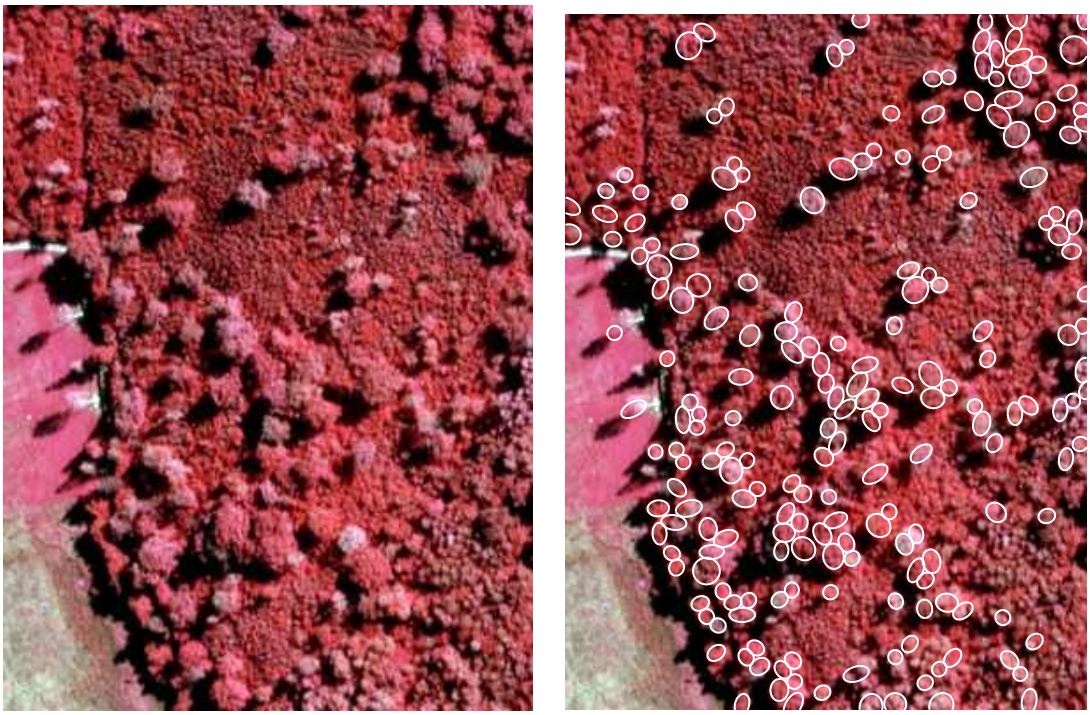

Figure 18: Left : 300×400 image of mixed height stands ©IFN. Right : extraction result, 5 millions of iterations ©INRIA.

4.4 Comments

By analysing the obtained results, we can see that both 2D and 3D models work correctly, and are quite robust w.r.t. the density in the stands. Some false alarms can be detected in the complete plantation image because of the data energy term which answers well for some objects (for example in the crops at the bottom of the image). However, these isolated objects could be deleted by a post-processing. Moreover, we remark that the border trees at the top left of this plantation are not detected, because of the deformation induced by the view angle. In this case, there is not enough shadow around the tree to extract it.

Generally speaking, the 2D and 3D algorithms give some useful information at the scale of the tree which is not conceivable without any help of image processing (it would take too much time for an operator to get these statistics, and also be very costly). For instance, the minimum mapped surface by IFN is 2.25 ha in France, while smaller surfaces like alignments or isolated trees are only statistically assessed via explorations and data measurements on the ground. To be able to respond to increasing (private or public) demands in precision, the inventories need to use such automatic algorithms. Once the extraction is performed, parameters such as the number of trees and their size, or the density of the stand, can be directly obtained by operators, and complete their knowledge on the forest. Finally, these results could interest plantation managers by giving them an estimation of the global wood volume of their stand.

5 Conclusion

In this paper we propose models to extract 2D and 3D vegetation resource parameters from high resolution aerial CIR photographs of forests. Based on a stochastic geometry approach, the models consider that each tree is one geometric object to be extracted in the image. It turns out to be efficient for segmenting the tree crowns in different kinds of stands, and gives precious information on the stand at the scale of the tree (number of trees, their size, density, ...). The 3D model for the data energy term is of particular interest because it gives access to statistics such as the wood volume. Such automatic algorithms are of economic importance because they can complete the work of forest managers and operators by giving them a more precise knowledge on the forest at a reasonable cost.

Future work will involve some tests on other kinds of images, and the implementation of a cooperation between the different models. Indeed, we could imagine a pre-processing of the image which would cut it into pieces and determine the correct model to be applied for vegetation resource extraction. Then, much work has to be carried out using shape and texture information, in order to try to associate a group of species or a class of age to one object by studying its shape, its texture, and the stand it is belonging to. The radiometry could then help to classify as much as possible the object (distinguishing between deciduous or conifer for instance).

6 Acknowledgments

The work of the first author has been partly supported by a grant from the MAS Laboratory of Ecole Centrale Paris. Part of this work has been conducted within the INRIA ARC "Mode de Vie" joint research program (Ariana and Digiplante Research Group - INRIA, MAS Laboratory - ECP, LIAMA - Chinese Academy of Sciences, IFN). The authors would like to thank particularly IFN for providing the data and for interesting discussions.

References

[1] H.E. Andersen. *Estimation of Critical Forest Structure Metrics through the Spatial Analysis of Airborne Laser Scanner Data*. PhD thesis, College of Forestry Resources, University of Washington, Seattle, USA, 2003.

[2] H.E. Andersen, S.E. Reutebuch, and G.F. Schreuder. Automated Individual Tree Measurement through Morphological Analysis of a LIDAR-based Canopy Surface Model. In *Proc. of the* 1^{st} *International Precision Forestry Symposium*, pages 11–21, June 2001.

[3] A. Baddeley and M.N.M. Van Lieshout. Stochastic Geometry Models in High-level Vision. In K.V. Mardia, editor, *Statistics and Images*, volume 1, pages 231–256. 1993.

[4] D.S. Culvenor, N. Coops, R. Preston, and K.G. Tolhurst. A Spatial Clustering Approach to Automated Tree Crown Delineation. In D.A. Hill and D.G. Leckie, editors, *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pages 67–80, Victoria, British Columbia, Canada, February 1998.

[5] X. Descombes, F. Kruggel, C. Lacoste, M. Ortner, G. Perrin, and J. Zerubia. Marked Point Process in Image Analysis : from Context to Geometry. In *SPPA Conference*, Castellon, Spain, April 2004. invited paper.

[6] M. Erikson. *Segmentation and Classification of Individual Tree Crowns*. PhD thesis, Swedish University of Agricultural Sciences, Uppsala, Sweden, 2004.

[7] F.A. Gougeon. Automatic Individual Tree Crown Delineation using a Valley-following Algorithm and Rule-based System. In D.A. Hill and D.G. Leckie, editors, *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pages 11–23, Victoria, British Columbia, Canada, February 1998.

[8] F.A. Gougeon, D.G. Leckie, I. Scott, and D. Paradine. Individual Tree Crown Species Recognition : the Nahmint Study. In D.A. Hill and D.G. Leckie, editors, *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pages 209–223, Victoria, British Columbia, Canada, February 1998.

[9] C. Lacoste. *Extraction de réseaux linéiques à partir d'images satellitaires et aériennes par processus ponctuels marqués.* PhD thesis, University of Nice-Sophia Antipolis, December 2004. (in French).

[10] M. Larsen. Individual Tree Top Position Estimation by Template Voting. In *Proc. of the Fourth International Airborne Remote Sensing Conference and Exhibition*, volume 2, pages 83–90, Ottawa, Ontario, June 1999.

[11] D.G. Leckie, C. Burnett, T. Nelson, C. Jay, N. Walsworth, F.A. Gougeon, and E. Cloney. Forest Parameter Extraction through Computer-based Analysis of High Resolution Imagery. In *Proc. of the Fourth International Airborne Remote Sensing Conference and Exhibition*, Ottawa, Ontario, June 1999.

[12] Inventaire Forestier National, editor. *Paysages de Forêts.* De Monza Publisher.

[13] M. Ortner. *Processus Ponctuels Marqués pour l'Extraction Automatique de Caricatures de Bâtiments à partir de Modèles Numériques d'Élévation.* PhD thesis, University of Nice-Sophia Antipolis, France, October 2004. (in French).

[14] A. Penttinen and D. Stoyan. Recent Applications of Point Process Methods in Forestry Statistics. *Statistical Science*, 15(1):61–78, 2000.

[15] G. Perrin, X. Descombes, and J. Zerubia. Adaptive Simulated Annealing for Energy Minimization Problem in a Marked Point Process Application. In *EMMCVPR Conf.*, Saint Augustine, Florida, USA, November 2005.

[16] G. Perrin, X. Descombes, and J. Zerubia. A Marked Point Process Model for Tree Crown Extraction in Plantations. In *ICIP Conf.*, University of Genoa, Italy, September 2005.

[17] G. Perrin, X. Descombes, and J. Zerubia. Point Processes in Forestry : an Application to Tree Crown Detection. Research Report 5544, INRIA, April 2005.

[18] A. Pinz. Tree Isolation and Species Classification. In D.A. Hill and D.G. Leckie, editors, *Proc. of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*, pages 127–139, Victoria, British Columbia, Canada, February 1998.

[19] R. Pollock. Individual Tree Recognition based on Synthetic Tree Crowns Image Model. In D.A. Hill and D.G. Leckie, editors, *Proc. of the International*

Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry, pages 25–34, Victoria, British Columbia, Canada, February 1998.

[20] H. Rue and A.R. Syversveen. Bayesian Object Recognition with Baddeley's Delta Loss. *Advances Applied Probability*, 30:64–84, 1998.

[21] M.N.M. van Lieshout. *Markov Point Processes and their Applications*. Imperial College Press, London, 2000.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399